



OICE

RAW SEQUENCE LISTING

DATE: 02/07/2002

PATENT APPLICATION: US/10/024,809

TIME: 18:03:27

Input Set : N:\Crf3\RULE60\10024809.txt

Output Set: N:\CRF3\02072002\J024809.raw

SEQUENCE LISTING

```

4 (1) GENERAL INFORMATION:
5     (i) APPLICANT: CFAIG, NANCY L
6     (ii) TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
7         ATP-DEPENDENT TRANSPOSITION PROTEINS
8     (iii) NUMBER OF SEQUENCES: 15
9     (iv) CORRESPONDENCE ADDRESS:
10        (A) ADDRESSEE: Anne Brown (Alston & Bird, LLP)
11        (B) STREET: 3605 Glenwood Ave.
12        (C) CITY: Raleigh
13        (D) STATE: NC
14        (E) COUNTRY: USA
15        (F) ZIP: 27608
16     (v) COMPUTER READABLE FORM:
17        (A) MEDIUM TYPE: Floppy disk
18        (B) COMPUTER: IBM PC compatible
19        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
20        (D) SOFTWARE: Patent In Release #1.0, Version #1.30
21     (vi) CURRENT APPLICATION DATA:
22        (A) APPLICATION NUMBER: US/10/024,809
23        (B) FILING DATE: 19-Dec-2001
24        (C) CLASSIFICATION:
25     (vii) PRIOR APPLICATION DATA:
26        (A) APPLICATION NUMBER: 09/027,169
27        (B) FILING DATE: 19-8-FEB-20
28     (viii) ATTORNEY/AGENT INFORMATION:
29        (A) NAME: Brown, Anne
30        (B) REGISTRATION NUMBER: 36,463
31        (C) REFERENCE/DOC# NUMBER: 5789-3
32     (ix) TELECOMMUNICATION INFORMATION:
33        (A) TELEPHONE: 919 420 2205
34        (B) TELEFAX: 919 881 3175
35 (2) INFORMATION FOR SEQ ID NO: 1:
36     (i) SEQUENCE CHARACTERISTICS:
37        (A) LENGTH: 1678 base pairs
38        (B) TYPE: nucleic acid
39        (C) STRANDEDNESS: single
40        (D) TOPOLOGY: linear
41     (ii) MOLECULE TYPE: DNA (genomic)
42     (ix) FEATURE:
43        (A) NAME/KEY: CDS
44        (B) LOCATION: 1..1668
45     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,809

DATE: 02/07/2002

TIME: 18:03:27

Input Set : N:\Crf3\RULE60\10024809.txt

Output Set: N:\CRF3\02072002\J024809.raw

65	ATG	AGT	GCT	ACC	CGG	ATT	CAA	GCA	GTT	TAT	CGT	GAT	ACG	GGG	GTA	GAG	48
66	Met	Ser	Ala	Thr	Arg	Ile	Gln	Ala	Val	Tyr	Arg	Asp	Thr	Gly	Val	Glu	
67	1				5				10					15			
68	GCT	TAT	CGT	GAT	AAT	CCI	TTT	ATC	GAG	GCC	TTA	CCA	CCA	TTA	CAA	GAG	96
69	Ala	Tyr	Arg	Asp	Asn	Pro	Phe	Ile	Glu	Ala	Leu	Pro	Pro	Leu	Gln	Glu	
70				20				25			30						
71	ACA	GTG	AAT	AGT	GCT	GCA	TCA	CTG	AAA	TCC	TCT	TTA	CAG	CTT	ACT	TCC	144
72	Ser	Val	Asn	Ser	Ala	Ala	Ser	Leu	Lys	Ser	Ser	Leu	Gln	Leu	Thr	Ser	
73			35					40				45					
74	TCG	GAC	TTC	CAA	AAG	TCC	CGT	GTT	ATC	AGA	GCT	CAT	ACC	ATT	TGT	CGT	192
75	Ser	Asp	Leu	Gln	Lys	Ser	Arg	Val	Ile	Arg	Ala	His	Thr	Ile	Cys	Arg	
76		50				55				60							
81	ATT	CCA	GAT	GAC	TAT	TTT	CAG	CCA	TTA	GGT	ACG	CAT	TTG	CTA	CTA	AGT	240
82	Ile	Pro	Asp	Asp	Tyr	Phe	Gln	Pro	Leu	Gly	Thr	His	Leu	Leu	Leu	Ser	
83	65				70				75				80				
85	CAG	CGT	ATT	TCC	GTC	ATG	ATT	CGA	GGT	GGC	TAC	GTA	GGC	AGA	AAT	CCT	288
86	Glu	Arg	Ile	Ser	Val	Met	Ile	Arg	Gly	Gly	Tyr	Val	Gly	Arg	Asn	Pro	
87				85				90			95						
89	AAA	ACA	GGA	GAT	TTA	CAA	AAG	CAT	TTA	CAA	AAT	GGT	TAT	GAG	CGT	GTT	336
90	Lys	Thr	Gly	Asp	Leu	Gln	Lys	His	Leu	Gln	Asn	Gly	Tyr	Glu	Arg	Val	
91			100					105			110						
93	CAA	ACG	GGA	GAG	TTG	GAG	ACA	TTT	CGC	TTT	GAG	GAG	GCA	CGA	TCT	ACG	384
94	Gln	Thr	Gly	Glu	Leu	Glu	Thr	Phe	Arg	Phe	Glu	Glu	Ala	Arg	Ser	Thr	
95		115					120				125						
97	GCA	CAA	AGC	TTA	TTG	TTA	ATT	GGT	TGT	TCT	GGT	AGT	GGG	AAG	ACG	ACC	432
98	Ala	Gln	Ser	Leu	Leu	Leu	Ile	Gly	Cys	Ser	Gly	Ser	Gly	Lys	Thr	Thr	
99		130				135			140								
101	TCT	CTT	CAT	CGT	ATT	CTA	GCC	ACG	TAT	CCT	CAG	GTG	ATT	TAC	CAT	CGT	480
102	Ser	Leu	His	Arg	Ile	Leu	Ala	Thr	Tyr	Pro	Gln	Val	Ile	Tyr	His	Arg	
103	145				150				155					160			
105	GAA	CTC	AAT	GTA	GAG	CAG	GIG	GIG	TAT	TTG	AAA	ATA	GAC	TCG	TCG	CAT	528
106	Glu	Leu	Asn	Val	Glu	Gln	Val	Val	Tyr	Leu	Lys	Ile	Asp	Cys	Ser	His	
107			165					170			175						
109	AAT	GGT	TCG	CTA	AAA	GAA	ATC	TGC	TTG	AAT	TTT	TTT	AGA	GCG	TTG	GAT	576
110	Asn	Gly	Ser	Leu	Lys	Gln	Ile	Cys	Leu	Asn	Phe	Phe	Arg	Ala	Leu	Asp	
111			180					185			190						
113	GGA	GCC	TTG	GGC	TGC	AAC	TAT	GAG	CGT	CGT	TAT	GGC	TTA	AAA	CGT	TAT	624
114	Arg	Ala	Leu	Gly	Ser	Asn	Tyr	Glu	Arg	Arg	Tyr	Gly	Leu	Lys	Arg	His	
115		195				200					205						
117	GGT	ATA	GAA	ACC	ATG	TTG	GCT	TTG	ATG	TCG	CAA	ATA	GCC	AAT	GCA	CAT	672
118	Gly	Ile	Glu	Thr	Met	Leu	Ala	Leu	Met	Ser	Gln	Ile	Ala	Asn	Ala	His	
119		210				215					220						
121	GCT	TTA	GGG	TTG	TTG	GTT	ATT	GAT	GAA	ATT	CAG	CAT	TTA	AGC	CGC	TCT	720
122	Ala	Leu	Gly	Leu	Leu	Val	Ile	Asp	Glu	Ile	Gln	His	Leu	Ser	Arg	Ser	
123	225				230				235				240				
125	CGT	TCG	GGT	GGA	TCT	CAA	GAG	ATG	CTG	AAC	TTT	TTT	GIG	ACG	ATG	GIG	768
126	Arg	Ser	Gly	Gly	Ser	Gln	Gln	Met	Leu	Asn	Phe	Phe	Val	Thr	Met	Val	
127			245					250			255						
129	AAT	ATT	ATT	GGC	GTA	CCA	GTG	ATG	TTG	ATT	GGT	ACC	CCT	AAA	GCA	GGA	816

RAW SEQUENCE LISTING

PATENT APPLICATION US/10/024,809

DATE: 02/07/2002

TIME: 18.03 27

Input Set : N:\Crf3\RULE60\10024809.txt

Output Set: N:\CRF3\02072002\J024809.raw

130	Asn	Ile	Ile	Gly	Val	Pro	Val	Met	Leu	Ile	Gly	Thr	Pro	Lys	Ala	Arg	
131				260					265					270			
133	GAG	ATT	TTT	GAG	GCT	GAT	TTC	CGG	TCT	GCA	CGT	AGA	GGG	GCA	GGG	TTT	864
134	Glu	Ile	Phe	Glu	Ala	Asp	Leu	Arg	Ser	Ala	Arg	Arg	Gly	Ala	Gly	Phe	
135			275					280					285				
137	GGA	GCT	ATA	TTG	TGG	GAT	CCT	ATA	CAA	CAA	ACG	CAA	CGT	GGA	AAG	CCC	912
138	Gly	Ala	Ile	Phe	Trp	Asp	Pro	Ile	Gln	Gln	Thr	Gln	Arg	Gly	Lys	Pro	
139		290					295						300				
141	AAT	CAA	GAG	TGG	ATC	GCT	TTT	ACG	GAT	AAT	CTT	TGG	CAA	TTA	CAG	CTT	960
142	Asn	Gln	Gln	Trp	Ile	Ala	Phe	Thr	Asp	Asn	Leu	Trp	Gln	Leu	Gln	Leu	
143	305					310					315				320		
145	TTA	CAA	CGC	AAA	GAT	GCG	CTG	TTA	TCG	GAT	GAG	GTC	CGT	GAT	GTG	TGG	1008
146	Leu	Gln	Arg	Lys	Asp	Ala	Leu	Leu	Ser	Asp	Glu	Val	Arg	Asp	Val	Trp	
147				325						330				335			
149	TAT	GAG	CTA	AGC	CAA	GGA	GTG	ATG	GAC	ATT	GTA	GTA	AAA	CTT	TTT	GTA	1056
150	Tyr	Glu	Leu	Ser	Gln	Gly	Val	Met	Asp	Ile	Val	Val	Lys	Leu	Phe	Val	
151			340					345					350				
153	CTC	GCT	CAG	CTC	CGT	GCG	CTA	GCT	TTA	GGC	AAT	GAG	CGT	ATT	ACC	GCT	1104
154	Leu	Ala	Gln	Leu	Arg	Ala	Leu	Ala	Leu	Gly	Asn	Glu	Arg	Ile	Thr	Ala	
155			355				360						365				
157	GGT	TTA	TTG	CGG	CAA	GTG	TAT	CAA	GAT	GAG	TTA	AAG	CGT	GTG	CAC	CCC	1152
158	Gly	Leu	Leu	Arg	Gln	Val	Tyr	Gln	Asp	Glu	Leu	Lys	Pro	Val	His	Pro	
159		370					375					380					
161	ATG	CTA	GAG	GCA	TTA	CGC	TCG	GGT	ATC	CCA	GAA	CGC	ATT	GCT	CGT	TAT	1200
162	Met	Leu	Glu	Ala	Leu	Arg	Ser	Gly	Ile	Pro	Glu	Arg	Ile	Ala	Arg	Tyr	
163	385				390					395			400				
165	TCT	GAT	CTA	GTC	GTT	CCC	GAG	ATT	GAT	AAA	CGG	TTA	ATC	CAA	CTT	CAG	1248
166	Ser	Asp	Leu	Val	Val	Pro	Gln	Ile	Asp	Lys	Arg	Leu	Ile	Gln	Leu	Gln	
167			405					410					415				
169	CTA	GAT	ATC	GCA	GCG	ATA	CAA	GAA	CAA	ACA	CCA	GAA	GAA	AAA	GGC	CTT	1296
170	Leu	Asp	Ile	Ala	Ala	Ile	Gln	Glu	Gln	Thr	Pro	Gln	Glu	Lys	Ala	Leu	
171			420					425					430				
173	CAA	GAG	TTA	GAT	ACC	GAA	GAT	CAG	CGT	CAT	TTA	TAT	CTG	ATG	CTG	AAA	1344
174	Gln	Glu	Leu	Asp	Thr	Glu	Asp	Gln	Arg	His	Leu	Tyr	Leu	Met	Leu	Lys	
175		435					440						445				
177	GAG	GAT	TAC	GAT	TCA	AGC	CTG	TTA	ATT	CCC	ACT	ATT	AAA	AAA	GCG	TTT	1392
178	Gln	Asp	Tyr	Asp	Ser	Ser	Leu	Ile	Pro	Thr	Ile	Lys	Lys	Ala	Phe		
179		450				455					460						
181	AGC	CAG	AAT	CCA	ACG	ATG	ACA	AGA	CAA	AAG	TTA	CTG	CGT	CTT	GTT	TTG	1440
182	Ser	Gln	Asn	Pro	Thr	Met	Thr	Arg	Gln	Lys	Leu	Leu	Pro	Leu	Val	Leu	
183	465				470					475			480				
185	CAG	TGG	TTG	ATG	GAA	GCG	GAA	ACG	GTA	GTG	TCA	GAA	CTA	GAA	AAG	CCC	1488
186	Gln	Trp	Leu	Met	Glu	Gly	Glu	Thr	Val	Val	Ser	Glu	Leu	Glu	Lys	Pro	
187			485					490					495				
189	TCC	AAG	AGT	AAA	AAG	GTT	TCG	GCT	ATA	AAG	GTA	GTG	AAG	CCC	AGC	GAG	1536
190	Ser	Lys	Ser	Lys	Lys	Val	Ser	Ala	Ile	Lys	Val	Val	Lys	Pro	Ser	Asp	
191			500					505					510				
193	TGG	GAT	AGC	TTG	CGT	GAT	ACG	GAT	TTA	CGT	TAT	ATC	TAT	TCA	CAA	CGC	1584
194	Trp	Asp	Ser	Leu	Pro	Asp	Thr	Asp	Leu	Arg	Tyr	Ile	Tyr	Ser	Gln	Arg	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,809

DATE: 02/07/2002

TIME: 18:03:27

Input Set : N:\Crf3\RULE60\10024809.txt

Output Set: N:\CRF3\02072002\J024809.raw

```

195          515          520          525
197 CAA CCT GAA AAA ACC ATG CAT GAA CGG TTA AAA GGG AAA GGG GTA ATA      1632
198 Gln Pro Glu Lys Thr Met His Glu Arg Leu Lys Gly Lys Gly Val Ile
199          530          535          540
201 GTG GAI ATG GCG AGC TTA TTT AAA CAA GCA GGT TAG CC      1670
202 Val Asp Met Ala Ser Leu Phe Lys Gln Ala Gly *
203 545          550          555

204 (2) INFORMATION FOR SEQ ID NO: 2:
205     (i) SEQUENCE CHARACTERISTICS:
206         (A) LENGTH: 555 amino acids
207         (B) TYPE: amino acid
208         (D) TOPOLOGY: linear
209     (ii) MOLECULE TYPE: protein
210     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
211 Met Ser Ala Thr Arg Ile Gln Ala Val Tyr Arg Asp Thr Gly Val Glu
212 1          5          10          15
213 Ala Tyr Arg Asp Asn Pro Phe Ile Glu Ala Leu Pro Pro Leu Gln Glu
214          20          25          30
215 Ser Val Asn Ser Ala Ala Ser Leu Lys Ser Ser Leu Gln Leu Thr Ser
216          35          40          45
217 Ser Asp Leu Gln Lys Ser Arg Val Ile Arg Ala His Thr Ile Cys Arg
218          50          55          60
219 Ile Pro Asp Asp Tyr Phe Gln Pro Leu Gly Thr His Leu Leu Leu Ser
220          65          70          75          80
221 Glu Arg Ile Ser Val Met Ile Arg Gly Gly Tyr Val Gly Arg Asn Pro
222          85          90          95
223 Lys Thr Gly Asp Leu Gln Lys His Leu Gln Asn Gly Tyr Glu Arg Val
224          100          105          110
225 Gln Thr Gly Glu Leu Glu Thr Phe Arg Phe Glu Glu Ala Arg Ser Thr
226          115          120          125
227 Ala Gln Ser Leu Leu Leu Ile Gly Cys Ser Gly Ser Gly Lys Thr Thr
228          130          135          140
229 Ser Leu His Arg Ile Leu Ala Thr Tyr Pro Gln Val Ile Tyr His Arg
230          145          150          155          160
231 Glu Leu Asn Val Glu Gln Val Val Tyr Leu Lys Ile Asp Cys Ser His
232          165          170          175
233 Asn Gly Ser Leu Lys Glu Ile Cys Leu Asn Phe Phe Arg Ala Leu Asp
234          180          185          190
235 Arg Ala Leu Gly Ser Asn Tyr Glu Arg Arg Tyr Gly Leu Lys Arg His
236          195          200          205
237 Gly Ile Glu Thr Met Leu Ala Leu Met Ser Gln Ile Ala Asn Ala His
238          210          215          220
239 Ala Leu Gly Leu Leu Val Ile Asp Glu Ile Gln His Leu Ser Arg Ser
240          225          230          235          240
241 Arg Ser Gly Gly Ser Gln Glu Met Leu Asn Phe Phe Val Thr Met Val
242          245          250          255
243 Asn Ile Ile Gly Val Pro Val Met Leu Ile Gly Thr Pro Lys Ala Arg
244          260          265          270
245 Glu Ile Phe Glu Ala Asp Leu Arg Ser Ala Arg Arg Gly Ala Gly Phe

```

RAW SEQUENCE LISTING

DATE: 02/07/2002

PATENT APPLICATION: US/10/024,809

TIME: 18:03:27

Input Set N:\Crf3\RULE60\10024809.txt

Output Set N:\CRF3\02072002\J024809.raw

```

269          275          280          285
271 Gly Ala Ile Phe Trp Asp Pro Ile Gln Gln Thr Gln Arg Gly Lys Pro
271          290          295          300
274 Asn Gln Glu Trp Ile Ala Phe Thr Asp Asn Leu Trp Gln Leu Gln Leu
271 305          310          315          320
277 Leu Gln Arg Lys Asp Ala Leu Leu Ser Asp Glu Val Arg Asp Val Trp
278          325          330          335
280 Tyr Glu Leu Ser Gln Gly Val Met Asp Ile Val Val Lys Leu Phe Val
281          340          345          350
283 Leu Ala Gln Leu Arg Ala Leu Ala Leu Gly Asn Glu Arg Ile Thr Ala
284          355          360          365
286 Gly Leu Leu Arg Gln Val Tyr Gln Asp Glu Leu Lys Pro Val His Pro
287          370          375          380
289 Met Leu Glu Ala Leu Arg Ser Gly Ile Pro Glu Arg Ile Ala Arg Tyr
290 385          390          395          400
292 Ser Asp Leu Val Val Pro Glu Ile Asp Lys Arg Leu Ile Gln Leu Gln
293          405          410          415
295 Leu Asp Ile Ala Ala Ile Gln Glu Gln Thr Pro Glu Glu Lys Ala Leu
296          420          425          430
298 Gln Glu Leu Asp Thr Glu Asp Gln Arg His Leu Tyr Leu Met Leu Lys
299          435          440          445
301 Glu Asp Tyr Asp Ser Ser Leu Leu Ile Pro Thr Ile Lys Lys Ala Phe
302          450          455          460
304 Ser Gln Asn Pro Thr Met Thr Arg Gln Lys Leu Leu Pro Leu Val Leu
305 465          470          475          480
307 Gln Trp Leu Met Glu Gly Glu Thr Val Val Ser Glu Leu Glu Lys Pro
308          485          490          495
310 Ser Lys Ser Lys Lys Val Ser Ala Ile Lys Val Val Lys Pro Ser Asp
311          500          505          510
313 Trp Asp Ser Leu Pro Asp Thr Asp Leu Arg Tyr Ile Tyr Ser Gln Arg
314          515          520          525
316 Gln Pro Glu Lys Thr Met His Glu Arg Leu Lys Gly Lys Gly Val Ile
317          530          535          540
319 Val Asp Met Ala Ser Leu Phe Lys Gln Ala Gly
320 545          550          555

```

322 (2) INFORMATION FOR SEQ ID NO: 3:

324 (i) SEQUENCE CHARACTERISTICS:

325 (A) LENGTH: 5926 base pairs

326 (B) TYPE: nucleic acid

327 (C) STRANDEDNESS: single

328 (D) TOPOLOGY: circular

330 (ii) MOLECULE TYPE: other nucleic acid

331 (A) DESCRIPTION: /desc = "pEM delta R.adj to 1"

336 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

338 TTATAGAGCAA TTGGGIGITA GTTTCAGCAA GCAAACATTA ACCATAGCTA ATGATTTATA      60
340 GCCAATATTA CCATTGGGGT ACCGAGCTCG AATTCATGG TCTGTTTCTT GTGIGAAATT      120
342 GTTATCCGCT CACAATTCCA CACATTATAC GAGTCGGATG ATTAAITGTC AACAGTCAT      180
344 TTCAGAAAT TGGCAGAAC CGTTAAGATG TCGGCGCAAA AAACATTATC CAGAAAGGGA      240
346 GTGCCTTGG AGCGACACGA ATTAAGCAGT GATTTACGAC CTGCACAGCC ATACCAACAGC      300

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/024,809

DATE: 02/07/2002

TIME: 18.03:28

Input Set : N:\Crf3\RULE60\10024809.txt

Output Set: N:\CRF3\02072002\J024809.raw

L:28 M:220 C: Keyword misspelled or invalid format. [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format. [(B) FILING DATE:]